

Adaptation and excitation effects underlying human auditory sensory memory

P. May^{1,2} and H. Tiitinen^{1,3}

¹*Department of Psychology, P.O. Box 13, 00014 University of Helsinki, Finland;* ²*Department of Mathematics, King's College London, WCR 2LS London, UK ;* ³*BioMag Laboratory, Medical Engineering Centre, Helsinki University Central Hospital, FIN-00029 HUS, Finland*

1 Introduction

The most prominent response in the auditory event-related field is the supratemporally generated N1m, which peaks about 100 ms after stimulus onset and lasts for approximately 100 ms. The N1m is exceptionally sensitive to stimulus rate, with its amplitude decreasing sharply as ISI is decreased. Previous research [1-3] has suggested that this attenuation is due to memory traces left by the previous stimuli. Further, as the duration of this memory trace coincides with estimates for the duration of auditory sensory memory, N1m could function as an index for this memory mechanism.

While the rate sensitivity of the N1m amplitude is well-established, the neural mechanisms underlying it remain unknown. Although the notion of "refractoriness" is sometimes used [see 1], the ERP literature suffers from a distinct opaqueness as to what cellular mechanism this might refer to. The cortex contains a range of slow inhibitory mechanisms [4] which could be considered natural candidates for the neural mechanisms of refractoriness and hence for N1m attenuation. These are spike frequency adaptation, related to non-synaptic after-hyperpolarizing currents lasting for seconds, and GABA_B mediated slow synaptic inhibition, lasting for hundreds of milliseconds.

In the present study, the theme of N1m and auditory sensory memory is further elaborated on. We model the generation of the N1m by using a computational model of cortical dynamics. The predictions of the model were tested with MEG measurements in humans.

2 Methods

2.1 Modelling methods

MEG is mostly due to synaptic currents in parallel apical dendrites of pyramidal cells oriented tangentially to the skull surface [5]. The N1m field pattern is roughly dipolar, implying a localized activation of a population of pyramidal cells. We modelled the neural dynamics of this population through an idealized, single-compartment

description of the pyramidal cell embedded in an approximation of the canonical cortical microcircuit [6, 7]. The population received recurrent excitation (rec), inhibition (inh) with fast and slow decay times, slow adaptation currents (adap), as well as excitatory (+) and inhibitory (−) thalamo-cortical afferent input (aff):

$$c_m \dot{u}(t) = I_{\text{leak}}(t) + I_{\text{rec}}(t) + I_{\text{inh}}^{(\text{fast})}(t) + I_{\text{inh}}^{(\text{slow})}(t) + I_{\text{adap}}(t) + I_{\text{aff}}^{(+)}(t) + I_{\text{aff}}^{(-)}(t), \quad (1)$$

where u is the membrane potential, I_x are the cross-membrane currents and c_m is the membrane capacitance. While the recurrent input was determined through u , the inhibitory input was received from a population of inhibitory cells, receiving excitatory input from the pyramidal cells as well as feedback inhibition. The leak current was determined by $I_{\text{leak}} = r_m^{-1}(E_{\text{leak}} - u)$, where E_{leak} is the resting membrane potential. The synaptic (rec, inh, aff) and adaptation (adap) currents were determined through respective channel reversal potentials E_x and conductances $G_x(t)$: $I_x(t) = [E_x - u(t)] G_x(t)$. For synaptic currents (rec, aff, and fast inh) the conductance changes were directly related to pre-synaptic firing. By assuming a uniform distribution of firing phases at M synapses, the total conductance can be approximated by $G(t) = M \bar{g} g_{\text{max}} f(t)$, where $\bar{g}$ is a normalizing constant, g_{max} is the peak conductance change at a single synapse, and f is the pre-synaptic firing rate. The latter was related to the membrane potential through $f(u) = 1[u] f_{\text{max}} [1 - \exp(u/s)]$, where f_{max} is the maximum firing rate, s is a constant, and $1[\cdot]$ is the step-function.

The adaptation currents depended on the firing rate of the pyramidal cells, and the associated conductance was described by the dynamical equation

$$\tau \dot{G}_a(t) = -G_a(t) + L f(u(t)), \quad (2)$$

where τ is the decay constant and L is a constant determining the strength of adaptation. For slow

inhibition, a similar equation was used to describe the conductance at the associated channel (with u being replaced by the membrane potential of the inhibitory cell population).

We derived an expression for the MEG in terms of the cross-membrane currents of the pyramidal cell population as:

$$B(t) = \frac{M_{\text{pop}} \mu_0 l}{8\pi r^2} [c_m \dot{u}(t) - 2I_i(t) + \Psi(l_{\text{leak}}(x, t))], \quad (3)$$

where M_{pop} is the number of cells in the population, l is the length of the cells, r is the distance between the population and the sensor, I_i is the sum of inhibitory currents (inh, adap), and the function Ψ is determined by the distribution of leak current along the cell (x). This approximation assumes that the excitatory input targets one end of the cell (the apical dendrites) and inhibitory input the other (the soma).

In simulations, we used 2-sec sequences of 50-ms stimuli with different ISIs. Parameter values were in the same range as those used in previous computational work [e.g., 8, 9]. We measured the pre-stimulus baseline amplitude (B_{base}) for the N th stimulus as the value of the MEG at stimulus presentation, as well as the peak amplitude (B_{peak}) of the response to the N th stimulus. The baseline-corrected peak amplitude was defined as $B_{\Delta p} = B_{\text{peak}} - B_{\text{base}}$.

2.2 Experimental methods

Auditory ERFs in 10 human subjects were measured for 2-sec binaurally presented stimulus sequences containing N tones, separated by 10-sec silences. The sequences contained 50-msec tones (including 10-ms rise and fall times, 70 dB SPL) presented with eight different ISIs ranging from 50 ($N = 40$) to 1950 ($N = 2$) msec. In addition, ERFs for a continuous 2-sec tone were recorded ($N = 1$). The subject was instructed to ignore the auditory stimuli and to concentrate on reading self-selected text.

MEG recordings were carried out in DC-mode with a 122-channel whole-head magnetometer using a sample rate of 400 Hz. Approximately 65 responses were averaged over a time period of 3.5 sec for each N . The responses were baseline-corrected with respect to a pre-sequence baseline of 100 ms and low-pass filtered at 30 Hz. Low frequency noise was eliminated by measuring the noise in the empty chamber and utilizing the signal-space projection method [10] to project its contribution out from the data. Eye movements were recorded from electrodes placed above the left eye and below the

outer canthus of the left eye. Artefacts were defined as amplitudes exceeding 150 μV .

Amplitude measurements were made for each N , each subject, and for both hemispheres. The pre-stimulus baseline amplitude B_{base} for each tone was determined by measuring the mean ERF in a 20 ms time window immediately preceding tone presentation. The peak amplitude B_{peak} for each tone was determined as the mean ERF in a 20 ms time window around the maximum of the response during a 150-ms post-onset time window. The baseline-corrected peak amplitude was defined as $B_{\Delta p} = B_{\text{peak}} - B_{\text{base}}$.

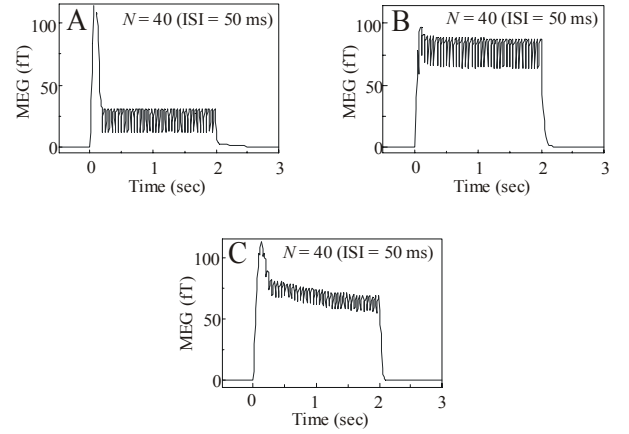


Figure 1: *Simulation results when $N = 40$ stimuli with an ISI of 50 ms were presented in a 2-sec window. Depending on the strength of the adaptation and inhibitory currents, the baseline-corrected peak amplitude of the response to the final stimulus (onset $t = 1.95$ sec) can be diminished for three reasons: In (A), the peak amplitude is attenuated. In (B), the peak amplitude is unaffected, but the pre-stimulus baseline amplitude is raised. In (C), a combination of attenuated peak amplitude and raised baseline amplitude contribute to a diminished baseline-corrected peak amplitude. The total number of cells was in the 10^4 - 10^5 range*

3 Results

3.1 Simulation results

The simulations demonstrated that an attenuation of the baseline-corrected peak amplitude for responses to stimuli presented at rapid rates can occur in three different scenarios. (1) When using parameter values allowing strong adaptation and inhibition to kick in after the presentation of the first stimulus, the peak amplitude of the responses to subsequent stimuli is strongly attenuated (Fig. 1A). This explains most of the attenuation of the baseline-

corrected peak amplitude, as the pre-stimulus baseline amplitudes are weakly affected. (2) When adaptation and inhibition is weakened, the peak amplitude for the response remains unaffected across the serial position of the stimulus (Fig. 1B). In addition, at the onset of each stimulus, excitation due to previous stimuli has decayed only minimally. This causes a strongly raised pre-stimulus baseline amplitude, and, consequently, a diminished baseline-corrected peak amplitude. (3) When intermediate strengths of adaptation and inhibition are used, the peak amplitude attenuates across serial position (Fig. 1C). Also, as depolarization due to previous stimuli is still present at stimulus presentation, the pre-stimulus baseline amplitude is raised for stimuli following the first. Therefore, the baseline-corrected peak amplitude is strongly attenuated, as in the previous two cases.

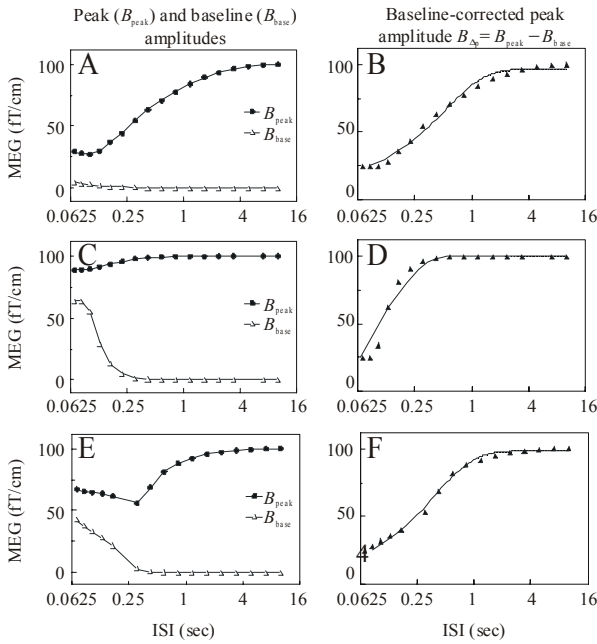


Figure 2: Simulation results showing the ISI-dependence of the pre-stimulus baseline amplitude (B_{base}), the peak amplitude (B_{peak}) and the baseline-corrected peak amplitude ($B_{\Delta p}$). Depending on the adaptation and inhibition, the behaviour of $B_{\Delta p}$ can be due to B_{peak} increasing (A&B) or B_{base} decreasing (C&D) as a function of ISI, or to a combination of these (E&F). In all cases $B_{\Delta p}$ is an increasing function of ISI which can be described by an exponentially saturating function used in [2,3].

The ISI-dependence of the baseline amplitude, the peak amplitude and the baseline-corrected peak

amplitude is demonstrated in Figure 2. Depending on the strength of inhibition and adaptation, the peak amplitude could either increase as a function of ISI (Figs. 2A & E), or remain relatively stable (Fig. 2C). Similarly, the baseline amplitude could either remain stable (Fig. 2A) or decrease as a function of ISI (Figs. 2C & E). However, in all cases, the baseline-corrected peak amplitude demonstrated strong stimulus rate dependence, growing monotonically as a function of ISI (Figs. 2B, D, F). Further, this behaviour could be characterized by an exponentially saturating function used in previous studies to describe N1m rate dependence [2, 3].

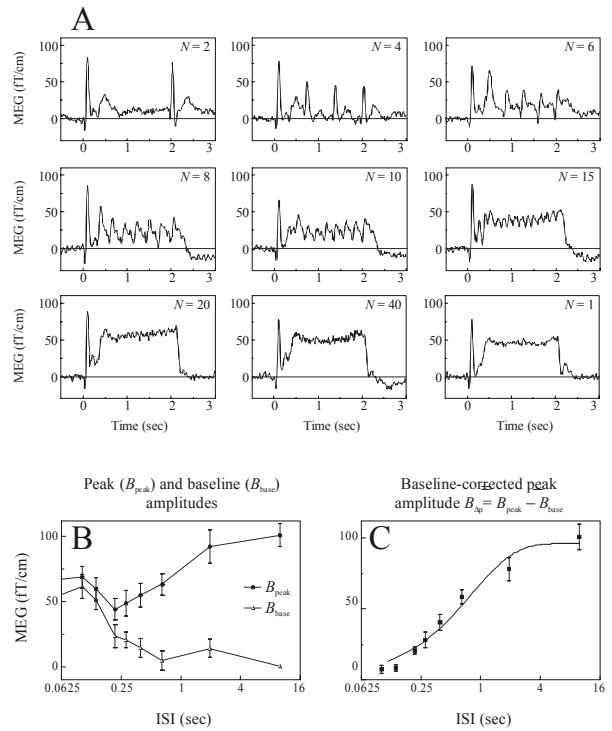


Figure 3: Experimental results averaged over 10 human subjects. In (A), magnetic measurements from the right fronto-temporal sensor displaying maximal N1m's. N refers to the number of stimuli presented in the time window 0 to 2 sec. In (B), the peak amplitude and the pre-stimulus baseline of the response to the last stimulus is plotted against the ISI of the sequence. Fast stimulation rates attenuate B_{peak} and raise B_{base} . In (C), the combined effect of this is the ISI-dependence of the baseline-corrected peak amplitude. An exponentially saturating function used in [2,3] has been fitted to the data.

3.2 Experimental results

The tones elicited prominent responses over both hemispheres, with the N1m fields being maximal over the fronto-temporal areas. In all conditions,

the first stimulus elicited a large N1m with a magnitude of around 100 fT/cm. For long ISIs (>280 ms, $N < 8$), the N1m's for the subsequent stimuli were attenuated to around 50 fT/cm. When shorter ISIs were used (<280 ms, $N > 8$), the peak amplitudes of the responses were again attenuated by 50%. In addition, the pre-stimulus baselines were raised, so that as the stimulus rate was increased, a sustained response became evident. The responses for the shortest ISIs (100 & 50 ms, $N = 20$ & 40) closely resembled the sustained field elicited by the continuous ($N = 1$) 2-sec tone. The measurements therefore revealed that with high stimulus rates (short ISIs, large N s) the baseline-corrected peak amplitude of the N1m is decreased by the combined effect of: (1) the pre-stimulus baseline amplitude being increased, and (2) the peak amplitude being decreased.

4 Discussion

The MEG recordings support our model that the N1m amplitude dependence on stimulus rate is due to (1) depolarization effects raising the pre-stimulus baseline and (2) adaptation and inhibition suppressing the peak amplitude of the N1m. The combined effect of these is the apparently strong attenuation of the response when pre-stimulus baseline correction is used. This correction has routinely been utilized in previous research [e.g., 1-3], which indicated that the N1(m) amplitude can be reduced by up to two orders of magnitude when rapid stimulus rates are employed. Our results contradict these findings in demonstrating that the relative variation in the peak amplitude of the auditory ERF across stimulation rates is due to amplitude reductions of only 50% (Fig. 3). Also, our observations show that the pre-stimulus baseline correction method can lead to serious under-estimations of cortical activity when fast stimulation rates are used.

Our combined computational and experimental results indicate that auditory sensory memory could manifest itself both in excitatory activity traces with a lifetime in the order of 100 ms (evident in the fast decaying sustained responses in Fig. 3A), and in inhibitory effects with lifetimes in the order of 1 sec (evident in the attenuation of the peak amplitude, Fig. 3A & B). These neural mechanisms might account for the short and long auditory memory stores suggested by behavioural evidence and experienced as sensation and memory, respectively [11].

Acknowledgements

We thank Mr. Ville Mäkinen for technical assistance. This work was supported by the Academy of Finland and the British Council.

References

1. R. Näätänen and T.W. Picton, "The N1 wave of the human electric and magnetic response to sound: a review and an analysis of the component structure", *Psychophysiology* **24**, 375-425, 1987.
2. Z.-L. Lu, S. J. Williamson, and L. Kaufman, "Human auditory primary and association cortex have differing lifetimes for activation traces", *Brain Research* **572**, 236-241, 1992.
3. Z.-L. Lu, S. J. Williamson, and L. Kaufman, "Behavioral Lifetime of Human Auditory Sensory Memory Predicted by Physiological Measures", *Science* **258**, 1668-1670, 1992.
4. G. M. Sheperd, *The Synaptic Organization of the Brain*, New York: Oxford University Press, 1990
5. M. Hämäläinen, R. Hari, R. J. Ilmoniemi, J. Knuutila, and O. V. Lounasmaa, "Magnetoencephalography - theory, instrumentation, and applications to noninvasive studies of the working human brain", *Reviews of Modern Physics* **65**, 413-497, 1993.
6. R. J. Douglas and K. A. C. Martin, "A functional microcircuit for cat visual cortex", *Journal of Physiology* **440**, 735-769, 1991.
7. R.J. Douglas, C. Koch, M. Mahowald, K.A.C. Martin, and H.H. Suarez, "Recurrent excitation in neocortical circuits", *Nature* **269**, 981-985, 1995.
8. M.A. Wilson and J.M. Bower, "The simulation of large-scale neural networks", in *Methods in Neural Modeling: From Synapses to Networks*, C. Koch and I. Segev, Eds. Cambridge MA: MIT Press, 1989, pp. 291-333.
9. C. Koch, M. Rapp, and I. Segev, "A brief history of time (constants)", *Cerebral Cortex* **6**, 93-101, 1996.
10. C.D. Tesche, M.A. Uusitalo, R.J. Ilmoniemi, M. Huottilainen, M. Kajola, and O. Salonen, "Signal-space projection of MEG data characterize both distributed and well-localized neuronal sources", *Electroencephalography and Clinical Neurophysiology* **95**, 189-200, 1995.
11. N. Cowan, "On short and long auditory stores", *Psychological Bulletin* **96**, 341-370, 1984.